

Density-functional theory study of gramicidin A ion channel geometry and electronic properties

Milica Todorović¹, D. R. Bowler^{2,3,4}, M. J. Gillan^{3,4}, Tsuyoshi Miyazaki¹

¹*National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan*

²*International Centre for Materials Nanoarchitectonics (MANA), National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan*

³*London Centre for Nanotechnology, UCL, 17-19 Gordon Street, London WC1H 0AH, UK*

⁴*Thomas Young Centre and Department of Physics & Astronomy, UCL, Gower Street, London WC1E 6BT, UK*

Understanding the mechanisms underlying ion channel function from the atomic-scale requires accurate ab initio modelling as well as careful experiments. Here, we present a density functional theory (DFT) study of the ion channel gramicidin A, whose inner pore conducts only monovalent cations and whose conductance has been shown to depend on the side chains of the amino acids in the channel. We investigate the ground-state geometry and electronic properties of the channel in vacuum, focusing on their dependence on the side chains of the amino acids. We find that the side chains affect the ground state geometry, while the electrostatic potential of the pore is independent of the side chains. This study is also in preparation for a full, linear scaling DFT study of gramicidin A in a lipid bilayer with surrounding water. We demonstrate that linear scaling DFT methods can accurately model the system with reasonable computational cost. Linear scaling DFT allows ab initio calculations with 10,000 to 100,000 atoms and beyond, and will be an important new tool for biomolecular simulations.

Key words: ion channel, gramicidin, DFT, linear scaling, electronic structure calculation.

1. Introduction

Ion channels reside in cell membranes and control the passage of ions, a vital function in regulating cell behaviour. The gramicidin A (gA) ion channel is one of the smallest and simplest ion channel systems[1–3], and was the target of the first antibiotic used in clinical practice. It is formed from a stack of two short beta-helix polypeptides, as shown in Fig. 1. Each polypeptide contains fifteen amino acids in a sequence of alternating left-handed and right-handed types. The latter helicity is rarely found in natural peptides, and owing to the alternating helicity, all peptide residues face outward into the lipid region, forming the unique helix structure with the beta type backbone structure. As is seen in Fig. 1, the channel structure forms a narrow pore, through which water molecules or ions can permeate in single file.

In spite of its simple structure, gA shows high selectivity towards ion permeation. It is selective for monovalent cations, such as H^+ , K^+ , Na^+ , or Cs^+ , showing no measurable permeability for anions or polyvalent cations[4]. Among many ion channels, gA was one of the first whose atomic resolution structure was available. Many experiments have been performed to investigate its transport properties and to clarify the mechanism of its high selectivity. Considering the existence of significant experimental information as well as its simple structure, the gA system is of great importance as a model membrane protein[3].

In the computational studies of the gramicidin-A, two high resolution structures have been commonly used so far. They are labelled as 1MAG and 1JNO in the PDB database, respectively. The former is obtained from solid-state NMR for the channel in DMPC membrane[5], while the latter one is from solution NMR, in detergent micelles[6]. While these two structures are topologically similar, they exhibit some variation in helix properties and residue Tryptophan 9 (Trp9) orientation (Fig. 2(a)). The 1MAG structure helix has $R=6.5$ residues per helix turn and $d=4.95$ Å helix turn height (pitch) in contrast to $R=6.3$ and $d=4.79$ Å of the 1JNO structure. The structural disparities could have originated from any aspect of experimental procedure, with little to indicate the optimal channel geometry: a classical potential MD study comparing the two structures in a lipid membrane environment predicted that the two backbones after dynamic relaxation starting from two structures become almost the same and that the Trp 9 spends 80 % of the time in the 1JNO orientation and 20 % in the 1MAG orientation[7]. Although we do not treat the ion permeation directly in this study, it is important to note the experimental result that the rate of the ion permeation is significantly changed by the replacement of the Tryptophan(Trp) residues, which have dipole moments ($\sim 2D$), with the other residues [6, 8–10]. When it is replaced with nonpolar Phenylalanine (Phe), the ion permeation rate is greatly reduced. The change of the ion permeation rate is larger if the number of substituted Trp's is larger, and also strongly depends on which Trp's are replaced; the mutation of Trp 9 is reported to show the largest change.

The system is an ideal target for theoretical studies. Although most theoretical studies were based on semi-microscopic models in the early stage, it is necessary to employ atomic-scale simulations, such as molecular dynamics (MD), for quantitative and detailed understanding of the permeation properties[11]. Many MD simulations of the gA ion channel system have been already reported[7, 12–16] and nowadays it is common to treat the gA system in the channel environment, that is gA with surrounding lipid bilayers (a simple model for the membrane) and bulk water. Such MD studies have clarified many aspects of the system, but they have a serious problem that the results sometimes depend on the type of the force field used in the calculations[15]. Since the gA system with its environment requires accurate modelling of many different types of interaction, the transferability of classical force fields may lead to poor description of specific interactions. For example, water molecules and ions are transported through the narrow pore of the gA channel. It is unreliable to use an unpolarisable force field for water molecules or ions in such a unique constrained space[17]. One of the well-known problem of the theoretical studies using the classical force fields is that the calculated energy barriers reported for the ion translocation in the gA pore are rather too high to explain the experimentally observed permeation rate[15, 18, 19]. It is therefore important to apply more accurate methods to the gA system.

Ab initio approaches based on density functional theory (DFT) have been playing important roles in clarifying the physical properties of various kinds of materials, though standard implementations of DFT are limited in the system size which they can model. A simplified gA model system has been modelled by DFT calculations to study the proton diffusion in the pore of the gA channel[20]. A recent quantum mechanics / molecular mechanics (QM/MM) study on the gA system showed the importance of polarisation of the water molecules in the gA pore[17, 21]. However, theoretical studies based on DFT are still very limited because of the demanding computational time and scaling with system size. Recent work has developed DFT methods which scale linearly with system size[22], allowing DFT calculations on systems with 10,000–100,000 atoms or more[23]. This brings the possibility to perform DFT calculations on entire biological molecules and their environment. In this work, we calculate the atomic and electronic structures of the isolated gA using the linear scaling DFT code CONQUEST[24, 25]. In particular, we investigate the effects of the side chains of the channel on its structural and

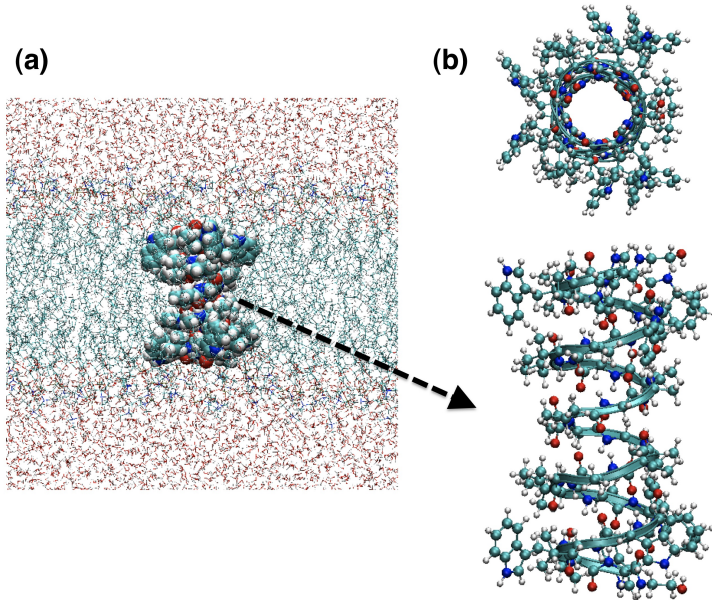


Figure 1: (a) Structure of gramicidin-A ion channel in lipid bilayers. The upper and lower area are bulk water region. (b) Top (upper) and side (lower) views of the gramicidin-A molecule. (Online version in colour.)

electronic properties. Although we only treat the isolated gA here, we still obtain important information about the structural stability and electronic structure of the gA molecule itself. We expect that such information would also help to clarify and improve the accuracy of a force field in the future. To our knowledge, it is the first DFT study to treat the whole gA molecule without any simplification. We will report linear scaling DFT calculations on the gA channel with full environment in a future publication.

The rest of the paper is organized as follows. In Sec. 2, we briefly explain the computational method used in this work. The results of our calculations on the atomic and electronic structures of the isolated gA molecule are presented in Sec. 3. Finally, concluding remarks are given in Sec. 4.

2. Computational Method: Local orbital and order-N methods

The calculations in this work are performed using the CONQUEST code, which is designed for very large-scale DFT calculations[23–26]. The method used in the code is based on density functional theory (DFT), with the pseudopotential technique. In this work, we employ the generalized gradient approximation (GGA) by Perdew, Burke and Ernzerhof (PBE)[27] for the exchange-correlation energy.

In CONQUEST we work with local orbitals $\{ \phi_{i\alpha}(r) \}$, which are centered on the position of each atom i with the index of orbital α . We represent the Kohn-Sham orbitals $\Psi_\nu(r)$ (ν : band

index) as linear-combinations of the local orbitals.

$$\Psi_\nu(r) = \sum_{i\alpha} c_{i\alpha}(\nu) \phi_{i\alpha}(r) \quad (2.1)$$

These local orbitals, which we call support functions, are themselves expressed by basis set functions. CONQUEST provides two types of basis set functions, B-splines on regular grids for accurate calculations since the basis set is systematically improvable[28], and pseudo atomic orbital (PAO) basis set for efficient calculations[29]. Users can choose one of the two basis set functions depending on the purpose. We only use the latter in this study.

Although CONQUEST can solve for the Kohn-Sham orbitals using a conventional $O(N^3)$ diagonalisation technique, the big advantage of using CONQUEST is that it can also employ $O(N)$ calculations. In the $O(N)$ calculations, the code uses the density matrix minimization method proposed by Li, Nunes and Vanderbilt (LNV)[30, 31]. Since the detailed explanation of the method has already been given in our previous papers, only the main points are explained in the following. In the density matrix minimization method, we calculate the density matrix instead of the Kohn-Sham orbitals. Formally, the density matrix is defined as

$$\rho(r, r') = \sum_{\nu} f_{\nu} \Psi_{\nu}^*(r) \Psi_{\nu}(r'). \quad (2.2)$$

The total energy within DFT can be expressed by the density matrix ρ and we optimise it instead of solving Kohn-Sham orbitals. In the LNV method, the matrix K is expressed as $K = 3LSL - 2LSLSL$ to keep the density matrix weakly idempotent. Then, we introduce a cutoff radius R_L for the off-diagonal elements of the L matrix in order to utilize the locality of density matrix. One important advantage of the LNV method is that it is variational with respect to R_L . We can examine the cutoff energy dependence of total energy easily and can evaluate an error from using a finite cutoff of R_L . This convergence behavior is reported in Sec. 3(c) for the DFT calculation of the isolated gA system.

In both diagonalisation and $O(N)$ methods, the code can perform non-selfconsistent (NSC) DFT calculations using the Harris-Foulkes energy functional with a charge density constructed from the superposition of atomic charge[32, 33]. Although the NSC method is less accurate than usual DFT calculations with the self-consistent field (SCF) method, it enables us to perform efficient structure optimization or molecular dynamics. Note that the code can calculate forces which are completely consistent with the NSC method[26, 34]. This method may be useful for our future study on gA systems, and its accuracy is also checked in this paper.

3. Results and Discussion

(a) Structural Properties

In this section, we report our results on the structural properties of the gA molecule in the vacuum. Here, we perform structure optimisation using two different high-resolution structures as initial atomic positions, 1MAG and 1JNO structures (Fig. 2(a)). In order to employ efficient structure optimisation, we first use the NSC method, then switch to the usual SCF method[34]. In both NSC and SCF methods, all atomic positions are relaxed using the standard conjugate gradient method until the maximum force component becomes smaller than 0.05 eV/Å. Note that the calculations in Sec. 3a and Sec. 3b are done by the standard diagonalisation method with the DZP basis set, explained in Sec. 2, and GGA-PBE functional[27]. We use only a Γ point

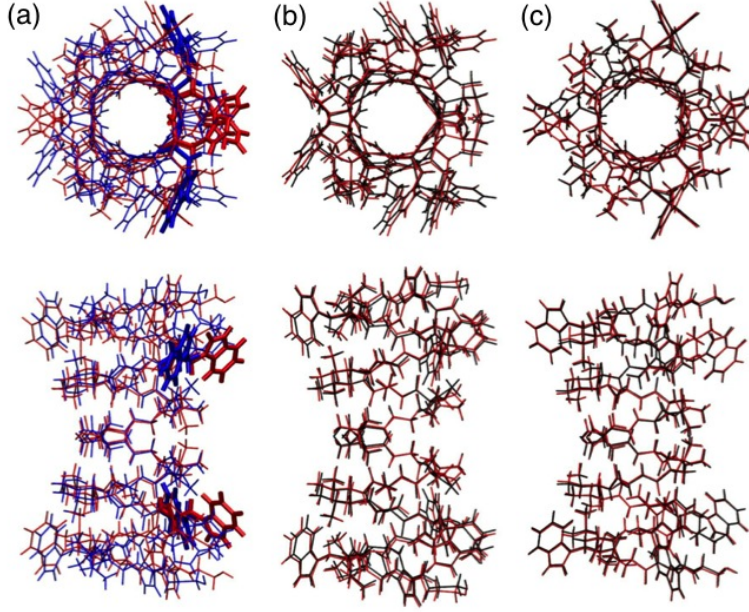


Figure 2: (a) Comparison of 1MAG (blue/dark) and 1JNO (red/light) experimental structures. *Ab initio* structural optimisation initial (black/dark) and final (red/light) structures in the (b) 1MAG case; (c) 1JNO case. (Online version in colour.)

for k -point sampling and the cutoff energy for the charge density grid is 150 Hartree. Periodic boundary is used with a unit cell of $23.3\text{\AA} \times 23.3\text{\AA} \times 35.0\text{\AA}$.

Figure 2(b) and (c) show the initial and optimised atomic positions starting from 1MAG and 1JNO structures, respectively. Although the original structures are obtained in different environment (in DMPC bilayers or in SDS micelles), we found that the optimised positions are close to the original ones in both cases. These figures also show that 1MAG has larger differences between the initial and final coordinates, especially for the backbone structure. In order to investigate the backbone structure in detail, we calculate the number of residues per helix turn R and the helix pitch d following the method in Ref.[35]. The root mean square deviations (RMSD) for the atoms of the backbone are also calculated. These structural parameters of 1MAG and 1JNO, before and after the structure optimisation, are listed in Table 1. The results show that the optimised 1MAG and 1JNO have different backbone structures. In addition, the change of the backbone structure is small in 1JNO case, while it is significant in 1MAG case, especially for helix pitch d . Although the original value of d in 1MAG case is already larger than that of 1JNO, the increase of d in 1MAG case is 0.15 \AA , which is larger than 0.08 \AA in 1JNO case. We observe a similar behaviour in the hydrogen bond lengths. The β -helix structure is stabilised by the hydrogen bonds formed between the backbone amide hydrogen and the carbonyl oxygen atoms parallel to the pore axis. The increase of the hydrogen bonds by the structure optimisation is 12% in 1MAG, while it is 3% in 1JNO case.

For the 1MAG structure, it is useful to know what causes the large change of the backbone structure during the optimisation: either the effect of backbone itself or that of side chains. One of the main differences between 1MAG and 1JNO structures is the different orientation of

Structure (SC)	RMSD	R^I	d^I	R^F	d^F	Δd
1MAG	0.29	6.5	4.95	6.55 (0.19)	5.10 (0.04)	0.14
1JNO	0.20	6.3	4.79	6.33 (0.07)	4.87 (0.06)	0.08
Alanine tube	0.21	6.5	4.95	6.52 (0.05)	4.99 (0.04)	0.04
Structure (NSC)	RMSD	R^I	d^I	R^F	d^F	Δd
1MAG	0.28	6.5	4.95	6.57 (0.05)	5.11 (0.04)	0.16
1JNO	0.14	6.3	4.79	6.33 (0.08)	4.84 (0.06)	0.05
Alanine tube	0.17	6.5	4.95	6.53 (0.05)	4.99 (0.03)	0.04

Table 1: RMSD between the initial (I) and optimised final (F) channel geometries and corresponding helix properties (number of residues per helix turn R and helix pitch d). Self consistent (upper) and non selfconsistent (lower) results are shown.

the indole ring of one Tryptophan (Trp9). The atomic positions of other side chains are also slightly different. These differences may cause the present different backbone structures. In order to investigate the effects of side chains on the backbone structure, we prepare a fictitious channel by replacing all amino acids of the gA with alanine, keeping the same backbone structure as the original 1MAG structure, shown in Fig. 3. Hereafter, we refer to this system as an alanine tube. We have relaxed the atomic positions of this system and analyse the values R and d of the resulting optimised structure. As is shown in Table 1, the change of the structure by the relaxation is found to be very small. The RMSD value is 0.21 Å and the change of d is only 0.04 Å. This result suggests that the large changes in the original 1MAG structure is due to the side chains, most likely steric hindrance effects. This explains the increase in the helix pitch during structure optimisation.

From the results so far, we can conclude that the original 1JNO structure is close to the optimised geometry in the gas phase, while 1MAG needs a large relaxation to reach the stable structure. By comparing the total energy, we find that initial 1JNO structure is more stable than 1MAG by 10.3 eV. The corresponding energy difference calculated by classical force fields is 10.4 eV by CHARMM (v27) and 19.7 eV by AMBER (v99)[36]. The value found using AMBER is larger than the DFT value, but these values are essentially consistent with the DFT result. Following structure optimisation with DFT, this large energy difference between the 1MAG and 1JNO structures is greatly reduced, but the final structure of 1JNO is still more stable than that of 1MAG by 1.8 eV. Judging from this result in the gas phase, 1JNO structure seems to be more appropriate for the stable geometry in a lipid bilayer. Since 1MAG geometry was measured in ordered bilayers, some fluctuations in side-chain orientation may have been overemphasized in the measurement. However, we should also comment that the energy difference between these structures may well become smaller in a lipid bilayer environment because the thickness of the lipid bilayers is usually larger than the length of gA. The strength of the interactions between gA and lipid molecules should be different depending on the length of the gA molecule and the 1MAG structure may be more stabilised by this effect than 1JNO, reducing the energy difference. We will address this in future work.

Next we report the stabilisation energy of the two helices of gA. When the gA ion channel is open for ions, the upper and lower helices must be associated to form a dimer. They are also sometimes dissociated by the lateral movement in the membrane (lipid bilayers) and in such cases the channel is closed for the ion permeation. Thus the stabilisation energy of the two helices can be regarded as a measure for the stability of the open channel structure. In this respect, the

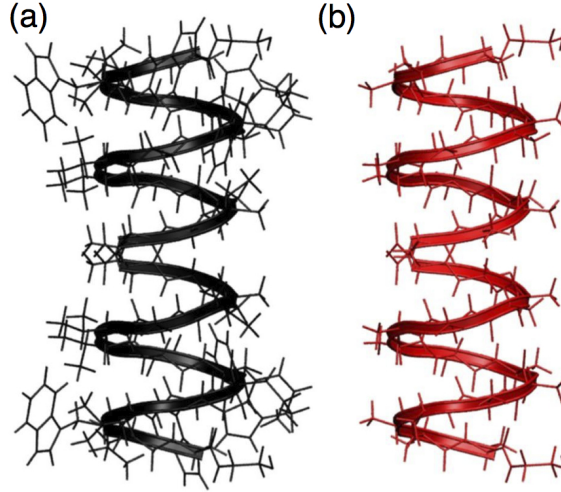


Figure 3: (a) Atomic structure of the gramicidin A (PDB code : 1MAG). (b) Alanine tube, which is made by replacing all the residues of gramicidin A by that of alanine. (Online version in colour.)

stabilisation energy is important and we calculate the energy defined as

$$E_s = E_{\text{dimer}} - 2E_{\text{mono}} \quad (3.1)$$

Here, E_{mono} is the total energy of the upper or lower peptides in isolation, whose atomic positions were relaxed. E_{dimer} is the total energy of the associated gA, calculated with the basis set superposition error (BSSE) correction using the counterpoise method[37]. With the optimised 1JNO structure, E_s is calculated to be 0.88 eV. This value is reasonable given the number of hydrogen bonds between the two helices, which have three N-H—O and three O-H—N bonds. Note that the exchange-correlation functional of PBE has been reported to be accurate for the expression of hydrogen bonds[38, 39]. In the membrane, this associated gA dimer structure will also be affected by the movement of surrounding lipid molecules. It will be important to compare this stabilisation energy with lipid-gA and lipid-lipid interactions in the future.

Finally, we note that, while we have mainly discussed results obtained with the SCF method, the differences between the NSC and SCF methods are small for the optimised structures (see Table 1). Since the NSC method is stable, and more efficient, it may play important roles in the future study on larger and more complicated full gA systems.

(b) Electronic Structure

In this section, we are particularly interested in the effects of the side chains of the ion channel on the electronic structure in the pore region of the gA. As mentioned in Sec. 1, it is known that the substitution of Trp significantly affects the rate of ion permeation[6, 8–10]. It is important to understand whether such differences come from the different electronic structure of the side chains or some other factors, for example, from the structural effects of the interactions of the side chains with lipid molecules. We note that there is a report showing that the conductance of a monovalent cation is correlated with the electrostatic interaction energy between the ion and the dipole moment of the indole ring of the tryptophans[9].

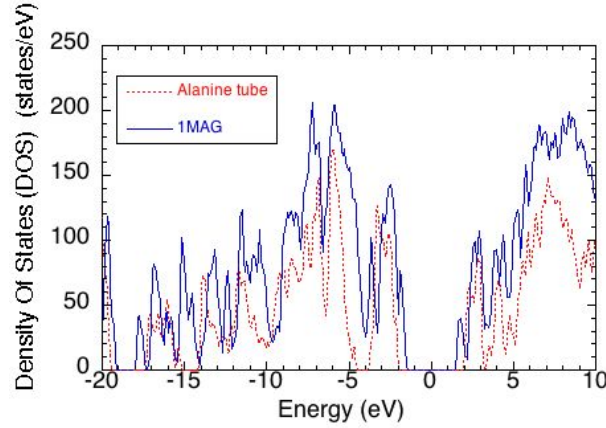


Figure 4: Electronic density of states for the alanine tube (red dotted line) and the 1MAG structure (blue solid line). (Online version in colour.)

First, we show in Fig. 4 the density of states of gA of 1MAG optimised structure (blue), as well as that of alanine tube (red) introduced in the last section. We also calculated the DOS of the 1JNO structure, but the difference to the 1MAG DOS is small, particularly around the band gap, so this is not shown for clarity. We can see that the density of states of the 1MAG shows a HOMO-LUMO energy gap of 3.3 eV. The energy gap of alanine tube is much larger than that of gA, at around 4.2 eV, implying that HOMO and/or LUMO of gA mainly come from its side chains. The alanine tube also shows a very low DOS around 4 eV below the Fermi level, which is not present in the gA DOS.

We studied the effects of side chains on the pore of the gA by investigating the charge density and the electrostatic potential. Ion transport depends critically on the electrostatic potential inside the channel pore, which in turn depends strongly on the charge density. The potential also affects the dynamics of the water molecules arranged in the pore. If the charge density is plotted as a function of distance from the channel axis, i.e. the centre of the pore, the two systems are found to be almost identical until the distance is greater than 4.4 Å, the approximate radius of the gA. We can thus conclude that the effects of the side chains on the charge distribution in the pore region are very small.

We next investigate the electrostatic potential, calculated as a sum of the local pseudopotential and the Hartree potential using the self-consistent charge density. Figure 5(a) shows a cross section of the potential perpendicular to the channel axis for the 1MAG structure, 8.8 Å along the axis away from the centre of the gA. In the figure, areas of the lower potential are located near the peptide bonds which form hydrogen bonds. In contrast, the region near the centre of the pore shows a high but flat potential, making it accessible to positively charged particles. Figure 5(b) shows the equivalent cross section for the alanine tube having a 1MAG backbone structure. There is a clear difference between Figs. 5(a) and (b) in the region of side chains, for example in the upper area. However, the two potentials are almost the same in the pore region. Figure 5(c) shows the difference of these two potential surfaces, implying that the two pores exhibit identical electrostatic potential within 0.01 eV. The same results are found at all points along the axis of the pore. We conclude that the effects of dipole moments of Trp's are well screened and the electronic structure of the side chains of the gA does not affect the electrostatic potential in its pore.

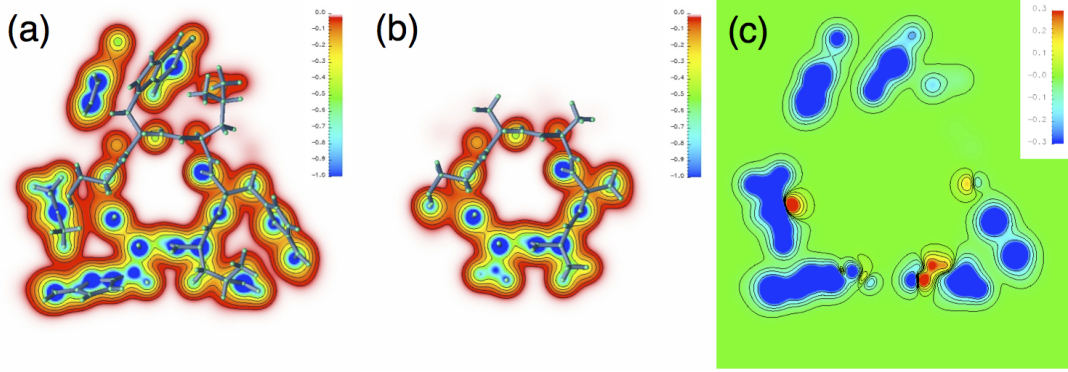


Figure 5: DFT electrostatic potential [eV]: (a) 1MAG geometry; (b) alanine channel; (c) 1MAG and alanine channel potential difference. Note the different scale in (c). (Online version in colour.)

(c) *Order- N calculations on the isolated gramicidin A*

It will be important in the future to perform all-atom DFT simulations of the gA channel in a lipid membrane with bulk water. To perform such large-scale DFT calculations, we will need to use an $O(N)$ method. As a preparation for these future studies, we have made preliminary $O(N)$ calculations on the isolated gA, in order to clarify the accuracy of the method on the gA systems. In particular, we can test the effect of the truncation of the density matrix since we can perform exact calculations (we have similarly tested convergence on a DNA ten-mer in water[39] which found essentially complete convergence for a range of 10\AA). Here, we use the NSC method because the convergence behaviour will be qualitatively the same for the SCF methods. For the basis sets, contracted DZP basis sets[29] are used.

Figure 6 shows the total energy calculated for different values of the cutoff, R_L . We find that all calculations are stable in the optimisation of L matrix. Since the present system contains only 552 atoms, we can employ the exact diagonalisation method and the total energy calculated by the method is also plotted by a dotted line in the figure. The inner panel of Fig. 6 shows the difference between the $O(N)$ and exact diagonalisation methods, on a log scale. This shows the total energy by the $O(N)$ method converges very fast towards the exact value. When the cutoff larger than 7.5\AA is used, the energy discrepancy is less than 1 meV per atom. This cutoff is small enough to realize actual calculations. In our previous $O(N)$ DFT study of a different system, Ge nano structures on Si(001) surface, we actually used 10.8\AA for the structure optimisation of the systems containing more than 20,000 atoms[40].

We expect that similar accuracy in the total energy would be obtained in the calculations of full gA systems, because the water and lipid molecules have large HOMO-LUMO energy gap, resulting in fast convergence of the total energy with respect to R_L . In fact, in our investigations of dry and hydrated DNA systems, we found that the difference caused by bulk water is extremely small[39]. Based on these results, we expect that the $O(N)$ calculations of the full gA systems will be both highly accurate and useful for understanding structure and function. These studies are already under way and will be presented in a future publication.

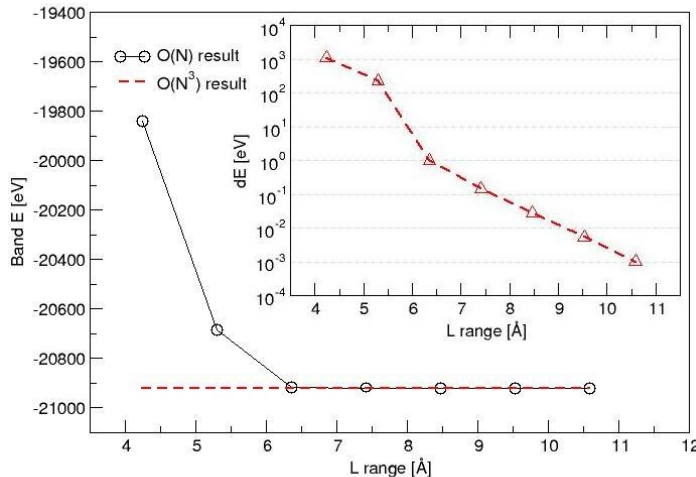


Figure 6: Order-N band energy behaviour with increasing range of auxiliary density matrix L (inset: convergence to diagonalisation result). (Online version in colour.)

4. Conclusions

In this work, we have performed a DFT study of gramicidin A (gA) to characterise the channel in isolation and examined its electronic structure. In particular, we have investigated the effects of the side chains on the structure and electronic properties of the channel.

For the structural stability of the gA in the gas phase, we have performed geometry optimisation starting from two different experimental atomic coordinates, 1MAG and 1JNO. We have found that the two calculations did not result in a common structure, but reinforced the difference in the helix properties, R and d . Although the original 1MAG structure has larger helix pitch than 1JNO, the optimisation of 1MAG structure exhibited a helix expansion. On the other hand, optimised 1JNO structure in the gas phase was found to be very close to the original geometry. To investigate the effects of side chains on the helix structure, we have introduced an artificial alanine tube system, which has a common backbone structure with 1MAG but all residues are replaced by that of alanine. We have found that the optimised helix pitch of the alanine tube is close to that of the original 1MAG structure. This result suggests that the original 1MAG structure in the gas phase includes steric hindrance between the side chains. Comparing the total energy of the optimised structures, 1JNO structure is more stable than 1MAG structure by 1.8 eV. However, this energy difference may become smaller in the membrane environment. Since the thickness of lipid bilayers is usually larger than the length of gA and 1MAG structure has smaller length mismatch, 1MAG structure may be more stabilized than 1JNO by the interactions between gA and lipid molecules. This aspect is interesting and should be clarified in the future study. In addition to the structure optimization of gA, we have also calculated the stabilisation energy for the upper and lower helices to make a dimer. This energy is important because the gA system shows a gating behaviour for the ion permeation by association

or dissociation of the two helices. The energy is calculated to be 0.8 eV, implying that reasonably strong hydrogen bonds are formed between the two helices.

We have also investigated the electronic structures of gA and the alanine tube. The charge density and electrostatic potential in the pore of the two systems have been compared to clarify the effects of side chains. The potential is relevant for the ion transport and the dynamics of the water molecules arranged in a single file in the channel. We found that the potential in the pore is almost identical between the two systems. The results imply that the sensitivity of the channel function to the type of residues cannot be explained by the electrostatic interactions made by the dipole of Trp's. This does not support some of the proposals in previous reports.

In the next step, it is desirable to simulate the system in the channel environment, that is gA in membrane with bulk water. We have already started to work on this large and complex gA system containing more than 16,000 atoms using an $O(N)$ DFT method. The code CONQUEST used in this work has an ability to treat very large systems containing many thousands of atoms, with the density matrix minimisation method. As a preparation for such large-scale DFT studies, we have applied the method to the isolated gA in this paper. By calculating the total energy dependence on the cutoff of the auxiliary density matrix R_L , we have found that the total energy converges very fast to the exact value. The errors from using the finite cutoff of R_L become negligibly small if we use R_L larger than 10 Å. It is encouraging and we believe that the method will become one of the standard theoretical techniques to study the atomic and electronic structures of complex biological systems.

Acknowledgment

The authors gratefully acknowledge useful discussions with Antonio Torralba, Michiaki Arita, Takao Otsuka and Wataru Shinoda. Most of the calculations in this work were done using the numerical materials simulator at the National Institute for Materials Science (NIMS), Japan. This work was partly supported by the Royal Society, Grant-in-Aid for Scientific Research on Innovative Areas (No. 22104005), the Strategic Programs for Innovative Research (SPIRE), MEXT, and the Computational Materials Science Initiative (CMSI), Japan.

References

- [1] Andersen, O. S., Koeppe II, R. E. & Roux, B., 2006 *Biological Membrane Ion Channels: Dynamics, Structure, and Applications*, chap. 2. Gramicidin Channels: Versatile Tools. Biological and Medical Physics, (Springer), 33–80.
- [2] Maffeo, C., Bhattacharya, S., Yoo, J., Wells, D. & Aksimentiev, A., 2012 Modeling and simulation of ion channels. *Chemical Reviews* **112**, 6250–6284. (doi:10.1021/cr3002609).
- [3] Kelkar, D. A. & Chattopadhyay, A., 2007 The gramicidin ion channel: A model membrane protein. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1768**, 2011–2025.
- [4] Myers, V. B. & Haydon, D., 1972 Ion transfer across lipid membranes in the presence of gramicidin A: ii. the ion selectivity. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **274**, 313 – 322. (doi:10.1016/0005-2736(72)90179-4).
- [5] Ketchum, R. R., Roux, B. & Cross, T. A., 1997 High-resolution polypeptide structure in a lamellar phase lipid environment from solid state NMR derived orientational constraints. *Structure* **5**, 1655–1669.

- [6] Townsley, L. E., Tucker, W. A., Sham, S. & Hinton, J. F., 2001 Structures of gramicidins A, B, and C incorporated into sodium dodecyl sulfate micelles. *Biochemistry* **40**, 11676–11686.
- [7] Allen, T. W., Andersen, O. S. & Roux, B., 2003 Structure of gramicidin A in a lipid bilayer environment determined using molecular dynamics simulations and solid-state NMR data. *J. Am. Chem. Soc.* **125**, 9868–9877.
- [8] Becker, M. D., Greathouse, D. V., Koeppe, R. E. & Andersen, O. S., 1991 Amino acid sequence modulation of gramicidin channel function: effects of tryptophan-to-phenylalanine substitutions on the single-channel conductance and duration. *Biochemistry* **30**, 8830–8839. (doi:10.1021/bi00100a015).
- [9] Hu, W. & Cross, T. A., 1995 Tryptophan hydrogen bonding and electric dipole moments: Functional roles in the gramicidin channel and implications for membrane proteins. *Biochemistry* **34**, 14147–14155. (doi:10.1021/bi00043a020).
- [10] Andersen, O. S., Greathouse, D. V., Providence, L. L., Becker, M. D. & Koeppe, R. E., 1998 Importance of tryptophan dipoles for protein function: 5-fluorination of tryptophans in gramicidin A channels. *Journal of the American Chemical Society* **120**, 5142–5146. (doi: 10.1021/ja980182l).
- [11] Edwards, S., Corry, B., Kuyucak, S. & Chung, S.-H., 2002 Continuum electrostatics fails to describe ion permeation in the gramicidin channel. *Biophys. J.* **83**, 1348–1360.
- [12] Woolf, T. B. & Roux, B., 1996 Structure, energetics and dynamics of lipid-protein interactions: A molecular dynamics study of the gramicidin A channel in a DMPC bilayer. *PROTEINS* **24**, 92–114.
- [13] Woolf, T. B. & Roux, B., 1994 Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer. *Proceedings of the National Academy of Sciences* **91**, 11631–11635.
- [14] Chiu, S.-W., Subramaniam, S. & E., J., 1999 Simulation study of a gramicidin/lipid bilayer system in excess water and lipid II. rates and mechanisms of water transport. *Biophys. J.* **76**, 1939–1950.
- [15] Allen, T. W., Baştuğ, T., Kuyucak, S. & Chung, S.-H., 2003 Gramicidin A channel as a test ground for molecular dynamics force fields. *Biophys. J.* **84**, 2159–2168.
- [16] Baştuğ, T., Gray-Weale, A., Patra, S. M. & Kuyucak, S., 2006 Role of protein flexibility in ion permeation: A case study in gramicidin a. *Biophysical journal* **90**, 2285–2296.
- [17] Bucher, D. & Kuyucak, S., 2009 Importance of water polarization for ion permeation in narrow pores. *Chemical Physics Letters* **477**, 207–210. (doi:http://dx.doi.org/10.1016/j.cplett.2009.06.069).
- [18] Baştuğ, T. & Kuyucak, S., 2006 Energetics of ion permeation, rejection, binding, and block in gramicidin A from free energy simulations. *Biophysical journal* **90**, 3941–3950.
- [19] Allen, T. W., Andersen, O. S. & Roux, B., 2006 Ion permeation through a narrow channel: Using gramicidin to ascertain all-atom molecular dynamics potential of mean force methodology and biomolecular force fields. *Biophys. J.* **90**, 3447–3468.
- [20] Sagnella, D. E., Laasonen, K. & Klein, M. L., 1996 Ab initio molecular dynamics study of proton transfer in a polyglycine analog of the ion channel gramicidin A. *Biophys. J* **71**, 1172–1178.

- [21] Timko, J. & Kuyucak, S., 2012 Investigation of polarization effects in the gramicidin a channel from ab initio molecular dynamics simulations. *The Journal of Chemical Physics* **137**, 205106. (doi:10.1063/1.4768247).
- [22] Bowler, D. R. & Miyazaki, T., 2012 $\mathcal{O}(N)$ methods in electronic structure calculations. *Rep. Prog. Phys.* **75**, 036503.
- [23] Bowler, D. R. & Miyazaki, T., 2010 Calculations on millions of atoms with density functional theory: linear scaling shows its potential. *J. Phys.: Condens. Matter* **22**, 074207. (doi: 10.1088/0953-8984/22/7/074207).
- [24] Bowler, D. R., Miyazaki, T. & Gillan, M. J., 2002 Recent progress in linear scaling *ab initio* electronic structure techniques. *J. Phys.:Condens. Matter* **14**, 2781–2798.
- [25] Bowler, D. R., Choudhury, R., Gillan, M. J. & Miyazaki, T., 2006 Recent progress with large-scale *ab initio* calculations: the CONQUEST code. *Phys. Stat. Sol. (b)* **243**, 989–1000.
- [26] Miyazaki, T., Bowler, D. R., Choudhury, R. & Gillan, M. J., 2004 Atomic force algorithms in density functional theory electronic-structure techniques based on local orbitals. *J. Chem. Phys.* **121**, 6186–6194.
- [27] Perdew, J. P., Burke, K. & Ernzerhof, M., 1996 Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868.
- [28] Hernández, E., Gillan, M. J. & Goringe, C. M., 1997 Basis functions for linear-scaling first-principles calculations. *Phys. Rev. B* **55**, 13485–13493.
- [29] Torralba, A. S., Todorović, M., V., Brázdová, Choudhury, R., Miyazaki, T., Gillan, M. J. & Bowler, D. R., 2008 Pseudo-atomic orbitals as basis sets for the $\mathcal{O}(N)$ DFT code CONQUEST. *J. Phys.:Condens. Matter* **20**, 294206.
- [30] McWeeny, R., 1960 Some recent advances in density matrix theory. *Rev. Mod. Phys.* **32**, 335–369. (doi:10.1103/RevModPhys.32.335).
- [31] Li, X.-P., Nunes, R. W. & Vanderbilt, D., 1993 Density-matrix electronic-structure method with linear system-size scaling. *Phys. Rev. B* **47**, 10891–10894.
- [32] Harris, J., 1985 Simplified method for calculating the energy of weakly interacting fragments. *Phys. Rev. B* **31**, 1770–1779.
- [33] Foulkes, W. M. C. & Haydock, R., 1989 Tight-binding models and density-functional theory. *Phys. Rev. B* **39**, 12520–12536.
- [34] Torralba, A. S., Bowler, D. R., Miyazaki, T. & Gillan, M. J., 2009 Non-self-consistent density-functional theory exchange-correlation forces for GGA functionals. *Journal of Chemical Theory and Computation* **5**, 1499–1505. (doi:10.1021/ct8005425).
- [35] Ketchem, R., Roux, B. & Cross, T., 1997 High-resolution polypeptide structure in a lamellar phase lipid environment from solid state NMR derived orientational constraints. *Structure (London, England : 1993)* **5**, 1655–1669.
- [36] Much larger difference between 1MAG and 1JNO is reported in Ref. [7].
- [37] Boys, S. & Bernardi, F., 1970 The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics* **19**, 553–566. (doi:10.1080/00268977000101561).

- [38] Ireta, J., Neugebauer, J. & Scheffler, M., 2004 On the accuracy of DFT for describing hydrogen bonds: dependence on the bond directionality. *The Journal of Physical Chemistry A* **108**, 5692–5698. (doi:10.1021/jp0377073).
- [39] Otsuka, T., Miyazaki, T., Ohno, T., Bowler, D. R. & Gillan, M. J., 2008 Accuracy of order-N density-functional theory calculations on DNA systems using CONQUEST. *J. Phys.:Condens. Matter* **20**, 294201.
- [40] Miyazaki, T., Bowler, D. R., Gillan, M. J. & Ohno, T., 2008 The energetics of hut-cluster self-assembly in Ge/Si(001) from linear-scaling DFT calculations. *J. Phys. Soc. Jpn.* **77**, 123706. (doi:10.1143/JPSJ.77.123706).